AMRL-TR-75-61

RESEARCH ON THE RECOGNITION AND ANALYSIS OF COMPLEX AND DYNAMIC IMAGERY

DEPARTMENT OF PSYCHOLOGY
MIAMI UNIVERSITY
OXFORD, OHIO 45056

OCTOBER 1975

Approved for public release; distribution unlimited

AEROSPACE MEDICAL RESEARCH LABORATORY
AEROSPACE MEDICAL DIVISION
Air Force Systems Command
Wright-Patterson Air Force Base, Ohio 45433

NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Organizations and individuals receiving announcements or reports via the Aerospace Medical Research Laboratory automatic mailing lists should submit the addressograph plate stamp on the report envelope or refer to the code number when corresponding about change of address or cancellation.

Do not return this copy. Retain or destroy.

Please do not request copies of this report from Aerospace Medical Research Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22151

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DODD 5230.0. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

HENNING E. VON GIERKE
Director
Biodynamics and Bionics Division
Aerospace Medical Research Laboratory

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: (18) AMRL (19) TR-75-61
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): (1) RESEARCH ON THE RECOGNITION AND ANALYSIS OF COMPLEX AND DYNAMIC IMAGERY.
5. TYPE OF REPORT & PERIOD COVERED: (9) Final report, 2 Jan 74 - 31 Aug 75
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): (10) Allan Pantle
8. CONTRACT OR GRANT NUMBER(s): (15) F33615-74-C-4032
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Miami University, Department of Psychology, Oxford, Ohio 45056
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: PROGRAM ELEMENT 62202F, PROJECT 7233, TASK 723305, WORK UNIT 72330511
11. CONTROLLING OFFICE NAME AND ADDRESS: Aerospace Medical Research Laboratory, Aerospace Medical Div, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio 45433
12. REPORT DATE: (11) Oct[redacted]975
13. NUMBER OF PAGES: (12) 63 p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (16) AF-7233 (17) 723305
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
18. SUPPLEMENTARY NOTES
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Vision; Human Visual System; Visual Information Processing; Spatial Frequencies; Harmonic Analysis of Visual Imagery; Independent Spatial Frequency Channels in Man; Visual Modulation Transfer Function; Transient Channels in Human Vision; Motion Perception; Temporal Frequency Response Functions; Visual Scanpaths; Visual Responses to Dynamic Imagery.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Recent literature on visual information processing contains considerable evidence that demonstrates the existence of functionally independent pattern (form) and transient (motion) systems in human vision. Research with complex static imagery and with dynamic imagery was conducted (and is reported herein) for the purpose of elucidating the operation of the pattern and transient systems, respectively.

Experiments on the perception of static, real-life scenes were con-

(Continued on reverse side.)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

3N

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20. Abstract (continued)

ducted within the framework of the Fourier-analyzer model, a model which states that the spatial frequency components of visual scenes are encoded in separate channels. Priorities for the acquisition of visual information in different spatial frequency ranges were studied in recognition experiments. Complementary research on eye scans was carried out in an effort to develop procedures for correlating eye fixations with the spatial frequency content of local regions of the scenes. The results suggest that relationships between recognition performance, eye scans, and the spatial frequency content of visual scenes can provide a useful basis for characterizing the search strategies of human observers in different kinds of perceptual tasks.

The transient system in human vision has functional properties which are different from those of the pattern system. The temporal frequency response function of the transient system was estimated with three different psychophysical techniques. Direct comparisons of the functions obtained with the different techniques suggest that there are two transient channels which were differentially involved in different tasks. The existence of the two channels was corroborated by a second series of experiments on the stroboscopic movement sensations elicited by the alternation of two multi-element stimulus frames. Knowledge of the response properties of the transient channels holds considerable promise for predicting human performance in detection, discrimination and tracking tasks with moving targets.

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Preface

The work discussed in this report was performed for the U.S. Air Force by Miami University (Oxford, Ohio) under the terms of contract No. F33615-74-C-4032, "Research on the Recognition and Analysis of Complex and Dynamic Imagery." This contract fell under Project 7233, "Applications of Biological Principles as Solutions to Air Force Needs in Signal Processing and Information Handling," Task 723305, "Automatic Pattern Recognition for Air Force Systems," Work Unit 72330511 (same title as contract). The work was performed during the period 2 January 1974 through 31 August 1975.

The contract monitors for this effort were Capt. J. W. Carl and Maj. R. Gagnon, Mathematics and Analysis Branch, Aerospace Medical Research Laboratory, Aerospace Medical Division (AFSC), Wright-Patterson Air Force Base, Ohio 45433.

Contents

List of Illustrations

Introduction and Summary

This report deals with spatiotemporal information processing in the human visual system. In an earlier report (Pantle, 1974), research on visual processing of static imagery was reviewed. A Fourier-analyzer model was used to summarize the research. Part of the new research described in that report involved the derivation of spatial filter functions (channel operating characteristics) which can be used to predict human performance in various types of perceptual tasks with static imagery. In the same report, preliminary evidence was summarized which suggested that the visual processing of dynamic imagery depends at least in part upon channels which are different from those involved in the processing of static imagery. Kulikowski and Tolhurst (1973) were the first investigators to draw a clear distinction between pattern (shape) channels and transient channels in human vision. They distinguished between two thresholds for temporally modulated sinusoidal gratings: (1) the minimum contrast at which flicker could be perceived (flicker threshold) and (2) the minimum contrast at which the spatial structure of a grating became distinct (pattern recognition threshold). The flicker and pattern recognition thresholds were independent functions of the spatial and temporal frequencies of the gratings, suggesting that the two types of thresholds reflect activity in two independent systems of channels. While a fair amount is known about the functional characteristics of pattern (spatial frequency) channels and their contributions to the perception of static imagery, very little is known about the functional

characteristics of transient or motion channels and their contributions to the perception of dynamic imagery.

In the present report original studies are described in which three different psychophysical methods were employed to isolate and describe the spatiotemporal response properties of transient or motion channels. In addition, new attempts to apply the Fourier-analysis model to the recognition and visual scanning of real-life scenes are reported.

Research on Spatiotemporal Information Processing by Transient (Motion) Channels in the Human Visual System: Studies with Sinusoidal Gratings

General Method

Subjects

The same three observers participated in all of the experiments on visual responses to dynamic imagery. Two of the observers (KLS and PLS) were university students, and the third observer (AJP) is the author of this report. Observer PLS was naive with respect to the purposes of the research. All observers were emmetropic or wore correcting lenses to obtain 20/20 acuity.

Apparatus

In all of the experiments on visual responses to dynamic imagery, the stimulus patterns were stationary sinusoidal gratings, drifting sinusoidal gratings, or spatiotemporal patterns formed by superimposing two sinusoidal drifting gratings. The patterns were displayed on the face of a Tektronix 5103 oscilloscope with a P31 phosphor. With minor modifications, the general technique described by Pantle (1973) was used to generate the stimulus gratings. In the motion aftereffect cancellation experiments described below, a spatiotemporal adapting pattern was employed, which was the sum of two oppositely drifting sinusoidal gratings. In order to produce this pattern, two audio-oscillators were placed in the external circuit, which supplied a complex modulating signal (the sum of two sinusoids) to the Z-axis of the

oscilloscope. The phase of one of the sinusoidal components of the modulating signal was continuously shifted in one direction relative to signal for the sweep generator of the scope, and the phase of the second component of the modulating signal was continuously shifted in the opposite direction. The observer viewed the display through a rectangular opening in an illuminated cardboard surround. The patterns completely filled the opening and subtended a visual angle of $4^{\circ}45'$ horizontally and $4^{\circ}06'$ vertically. The cardboard surround was circular and its outside diameter was $10^{\circ}53'$. The luminance and color of the surround were matched to the space-average luminance (1.8 mL) and color (yellow-green) of the oscilloscope patterns. The observer viewed the display binocularly with natural pupils from a distance of 107 cm. A small black fixation spot was located in the center of the stimulus pattern.

Experiment I: Threshold Flicker Sensitivity

In the first experiment we obtained measurements of flicker or movement thresholds in the manner described by Kulikowski and Tolhurst (1973). These measurements were made so that the temporal frequency response functions obtained by the other psychophysical methods described below could be directly compared with flicker threshold data from the same observers. Each observer adjusted the contrast of a 1-c/deg, drifting sinusoidal grating until he could just detect the movement of the grating bars or the flicker produced by the drifting grating, whichever required the least contrast. Five contrast threshold measurements were obtained at each of nine drift rates (temporal

frequencies) between 0.3 and 20 Hz. One measurement at each of the nine drift rates constituted a block, and the order of presentation of the drift rates was random within each block.

Contrast thresholds (log percent contrast) as a function of the drift rate (temporal frequency) of the grating are shown in Fig. 1. Each function represents the data for a different observer. Each data point is the mean of the five threshold estimates obtained for each observer at a given drift rate. The standard errors of the means plotted in Fig. 1 ranged from 0.01 to 0.06 with a mean of 0.03 (about 7%). As reported in previous studies, the contrast threshold for flicker or movement is lowest at an intermediate temporal frequency (0.4-0.6 log Hz--2.5-5 Hz) and increases at both higher and lower temporal frequencies.

Experiment II: Cancellation Technique for Measuring Motion Aftereffect

It has been suggested (e.g., Kulikowski and Tolhurst, 1973) that the temporal frequency response function defined by flicker thresholds reflects the activity of transient or movement-sensitive channels that operate independently of and in parallel with shape or spatial frequency channels. The definition of the transient channels has relied upon the ability of an observer to maintain a distinction between two threshold criteria: a flicker threshold criterion and a pattern threshold criterion. Moreover, the evidence for separate transient channels has only been obtained in experiments with threshold-level signals. In Experiment II we used a psychophysical cancellation technique which was

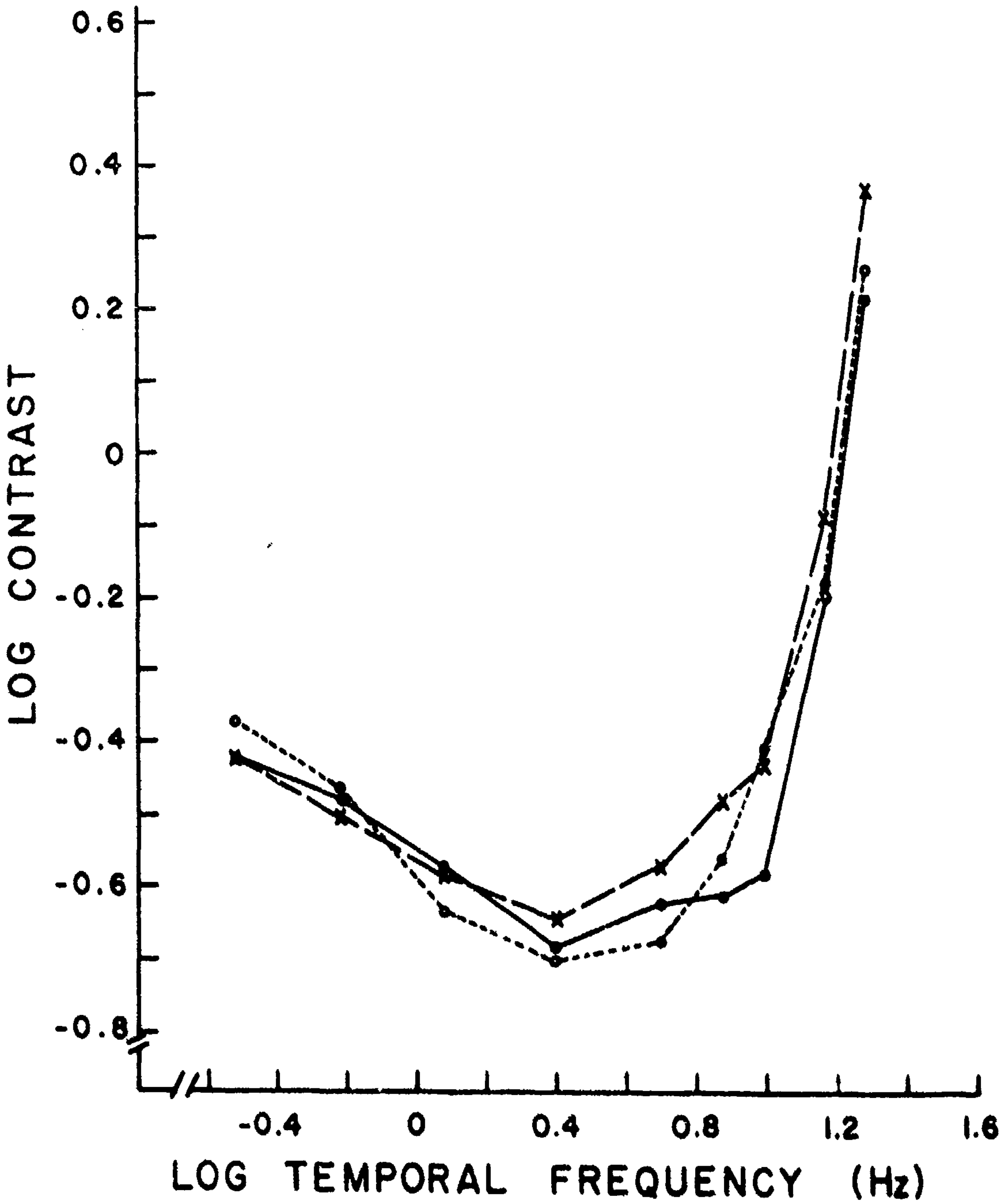

Fig. 1. Contrast thresholds for detecting the movement or flicker of drifting sinusoidal gratings as a function of their drift rate. Observer PLS: ●s; observer AJP: Xs; and observer KLS: Os.

designed to isolate the responses of transient or motion-sensitive channels and which required an observer to act only as a null instrument. In addition, it allowed us to study the response of transient channels or movement-sensitive channels to suprathreshold patterns.

The cancellation technique makes use of the motion aftereffect phenomenon--after an observer views contours moving uniformly in one direction for a prolonged period of time, stationary contours will appear to move in the opposite direction. In Experiment II, a motion aftereffect was generated by having observers adapt to a 1-c/deg sinusoidal grating which moved rightward (hereafter this grating will be referred to as $\underline{f}_R$). A stationary sinusoidal test grating (1 c/deg) that was presented afterwards appeared to move to the left. The adapting and test patterns alternated, 8 sec for the adapting stimulus and 2 sec for the test stimulus. After the first few adapting-test cycles (trials), a second sinusoidal grating (also 1 c/deg) was superimposed upon $\underline{f}_R$ during the adaptation period. The second adapting grating (hereafter called $\underline{f}_L$) moved to the left. By adjusting the contrast of $\underline{f}_L$, an observer could cancel or null the motion aftereffect produced by $\underline{f}_R$. For low contrasts of $\underline{f}_L$, the aftereffect generated by $\underline{f}_R$ was weakened; at high contrasts, the aftereffect generated by $\underline{f}_R$ was overridden and reversed; at some intermediate contrast (hereafter called the null contrast), the aftereffect generated by $\underline{f}_R$ was just cancelled (i.e., the test pattern appeared stationary).

The determination of the null contrast required a series of

adjustments over a number of trials. Initially the contrast of $\underline{f}_L$ was set well below the null contrast at a randomly determined starting point. Thereafter the observer increased or decreased the contrast of $\underline{f}_L$ depending upon the direction and magnitude of the aftereffect he saw on each trial. When the contrast of $\underline{f}_L$ was in the vicinity of the null contrast, it was necessary for the observer to see the aftereffects on a number of trials before making further changes in the contrast of $\underline{f}_L$. At each contrast level of $\underline{f}_L$, a few trials were required for the adaptation process underlying the aftereffect to reach a steady state. The observer terminated the adjustment series when he was satisfied that he saw no aftereffect during the test period. The contrast of $\underline{f}_L$ at that time provided one estimate of the null contrast. Each adjustment series typically lasted 10-15 min.

Three independent estimates of the null contrast were obtained at each of nine different drift rates (temporal frequencies) of $\underline{f}_L$. The nine drift rates were arranged in blocks, and the order of presentation within a block was random. Variations of the null contrast as a function of the drift rate of $\underline{f}_L$, for a fixed contrast and drift rate of $\underline{f}_R$, define a suprathreshold temporal frequency response function for motion-sensitive channels. The function is a constant-response function since it describes changes of input (null contrast) required to produce a constant output, i.e., a constant amount of adaptation produced by $\underline{f}_L$ to cancel a standard (fixed) directional aftereffect produced by $\underline{f}_R$.

A temporal frequency response function was obtained for each

of three drift rates of $\underline{f}_R$. For each function the drift rate of $\underline{f}_R$ was held constant at either 1.25, 5, or 15 Hz. At each drift rate the contrast of $\underline{f}_R$ was also held constant at one of the values given in Table 1. The physical contrasts were approximately 10 times the observers' flicker thresholds. The actual factor is given in parentheses behind the physical contrast. The test grating contrast used for each observer is also given in Table 1. The test contrasts were approximately 10 times the observers' pattern thresholds for the test grating.

Table 1

Contrast values used for each of a number of different rates of $\underline{f}_R$ for each subject to obtain a temporal frequency response function

Observer	Drift rate of $\underline{f}_R$ 1.25	5	15	Test grating
PLS	0.23 (11X)	0.26 (11X)	0.62 (11X)	0.44 (10X)
AJP	0.26 (11X)	0.19 (7X)	0.51 (7X)	0.46 (10X)
KLS	0.21 (11X)	0.23 (11X)	0.73 (11X)	0.59 (10X)

The results of Experiment II are shown in Figs. 2-4 for observers PLS, AJP, and KLS, respectively. The abscissa in each figure is the temporal drift rate of $\underline{f}_L$. The ordinate is the logarithm of the ratio of the contrast of $\underline{f}_L$ to $\underline{f}_R$. If the null contrast of $\underline{f}_L$ were equal to the contrast of $\underline{f}_R$, then the logarithm of the relative contrast would be equal to 0.0 and would lie on the dashed line. If the contrast of $\underline{f}_L$ had to be

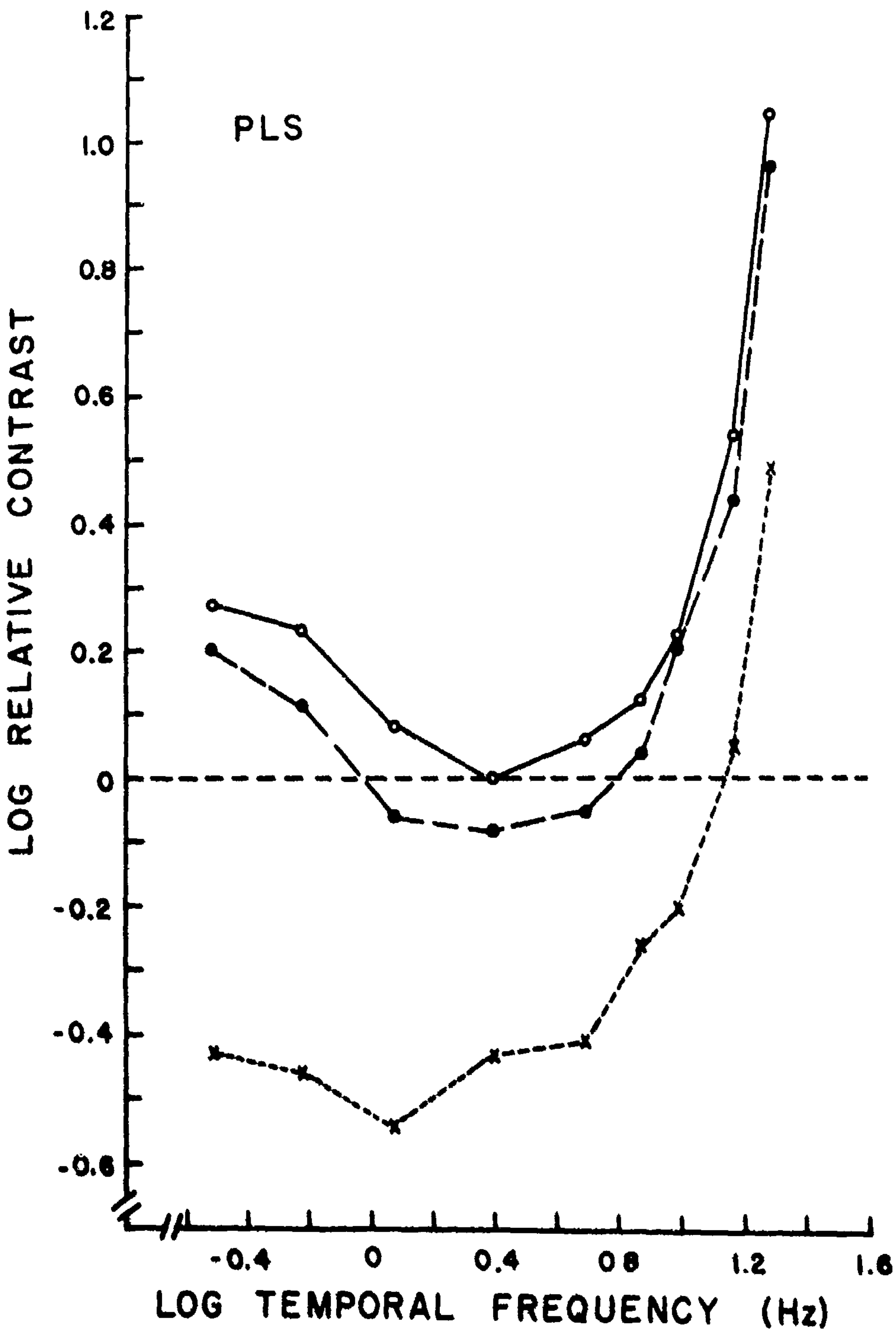

Fig. 2. Relative null contrast required for one adapting grating ($\underline{f}_L$) to cancel the aftereffect generated by a second adapting grating ($\underline{f}_R$) as a function of the drift rate (temporal frequency) of $\underline{f}_L$. Data depicted by closed circles, open circles, and Xs were obtained with $\underline{f}_R$ drift rates of 1.25, 5, and 15 Hz, respectively. Observer PLS.

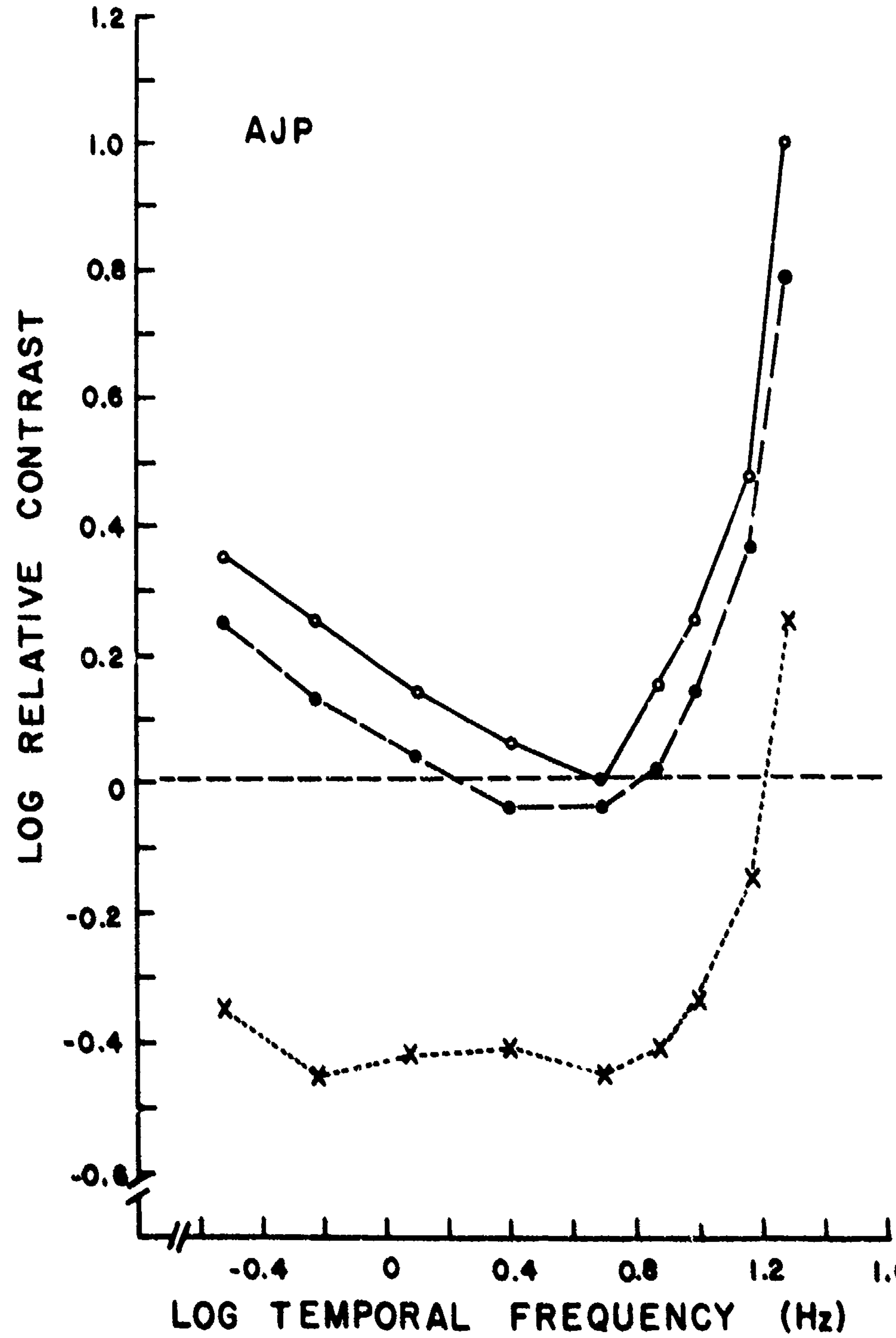

Fig. 3. Relative null contrast required for one adapting grating ($\underline{f}_L$) to cancel the aftereffect generated by a second adapting grating ($\underline{f}_R$) as a function of the drift rate (temporal frequency) of $\underline{f}_L$. Data depicted by closed circles, open circles, and Xs were obtained with $\underline{f}_R$ drift rates of 1.25, 5, and 15 Hz, respectively. Observer AJP.

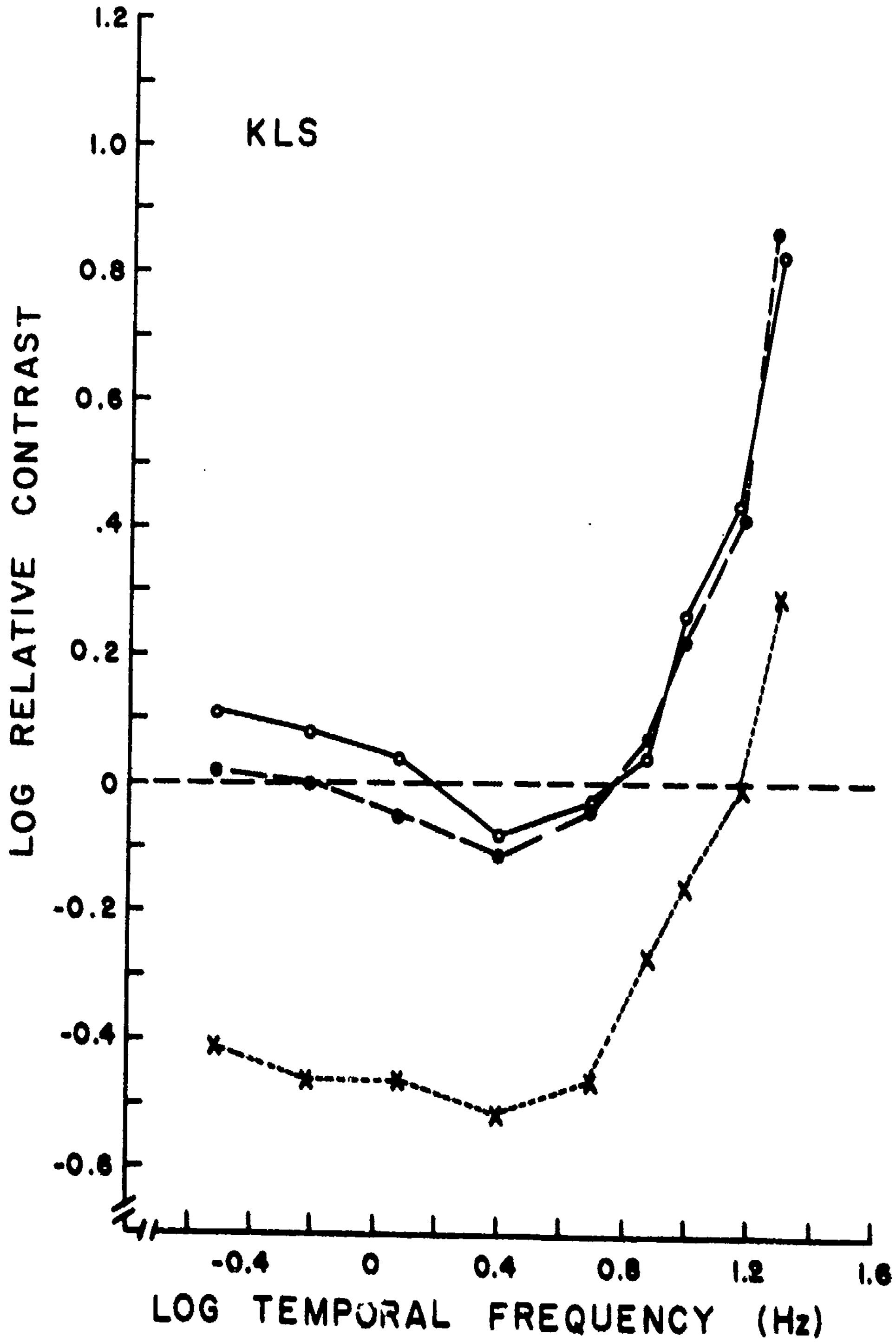

Fig. 4. Relative null contrast required for one adapting grating (f_L) to cancel the aftereffect generated by a second adapting grating (f_R) as a function of the drift rate (temporal frequency) of f_L. Data depicted by closed circles, open circles, and Xs were obtained with f_R drift rates of 1.25, 5, and 15 Hz, respectively. Observer KLS.

greater than that of $\underline{f}_R$ in order to cancel the aftereffect generated by $\underline{f}_R$, then the data point would fall above the dashed line. If the null contrast of $\underline{f}_L$ were less than that of $\underline{f}_R$, then the data point would fall below the dashed line. Each data point is the mean of the three estimates of the null contrast obtained for each condition. The standard errors of the means ranged from 0.00 to 0.08 and had a mean of 0.035 (8%). Each function represents the data obtained with a different drift rate of $\underline{f}_R$. The functions for the three observers are essentially the same. In order to cancel the aftereffect generated by $\underline{f}_R$ when it drifted at 1.25 Hz (function drawn with solid circles), the observer adjusted the contrast of $\underline{f}_L$ when it drifted at 1.25 Hz (0.1 log Hz) to approximately the same value (within experimental error) as that of $\underline{f}_R$. (For each observer the data point lies close to the dashed line.) The relative null contrast increased for lower temporal frequencies of $\underline{f}_L$ (< 1.25 Hz) and for high temporal frequencies of $\underline{f}_L$ (> 5 Hz), and it was at a minimum in the vicinity of 2.5-5 Hz (log 0.4-0.7 Hz). This means that the aftereffect generated by $\underline{f}_R$ was most readily cancelled by $\underline{f}_L$ when $\underline{f}_L$ moved at a drift rate of approximately 2.5 to 5 Hz. Other temporal frequencies were less effective. The same pattern of results was obtained when the drift rate of $\underline{f}_R$ was 5 Hz (function drawn with open circles). The relative null contrast was minimal at 2.5-5 Hz (log 0.4-0.7 Hz) and increased at both higher and lower temporal frequencies. The temporal frequency response functions for the 1.25- and 5-Hz drift rates of $\underline{f}_R$ are remarkably similar to the flicker threshold functions shown in Fig. 1 for the same observers.

The temporal frequency response function obtained with an $\underline{f}_R$ drift rate of 15 Hz had a low-pass characteristic; i.e., the least null contrast was required at low temporal frequencies of $\underline{f}_L$ and in general it increased with the temporal frequency of $\underline{f}_L$.

Experiment III: Velocity Difference Thresholds

Experiments I and II dealt with the temporal contrast sensitivity of motion channels which underlie (1) the detection of temporal intensity modulations at threshold and (2) the motion aftereffect. In Experiment III we determined the sensitivity of the visual system to changes in the drift rate of sinusoidal gratings. In particular we attempted to determine (1) whether sensitivity to small changes of drift rate ($\Delta \overline{DR}$) would vary as a function of drift rate ($\overline{DR}$) and, if so, (2) whether the relationship between $\Delta\overline{DR}$ and $\overline{DR}$ is the same as the obtained relationship between temporal contrast sensitivity and drift rate of sinusoidal gratings. If the same motion channels underlie both sensitivity to temporal modulations and to changes of drift rate, then it seems reasonable to expect that sensitivity to changes of drift rate ought to be greatest at temporal frequencies where flicker sensitivity is greatest. In other words, sensitivity to flicker and to changes of flicker rate ought to covary as a function of the temporal frequency of drifting gratings.

The method of constant stimuli was used to measure the ability of an observer to discriminate differences of drift rates of sinusoidal gratings. A trial consisted of two discrete temporal intervals. A grating moving rightward at one of five preselected

test velocities was presented in the first interval. A grating moving leftward at a standard velocity was presented in the second interval. Two of the test velocities were faster, two were slower, and one was equal to the standard velocity. In order to avoid sharp temporal transients when the gratings were switched "on" and "off," the contrast of the grating presented in each interval was increased linearly from 0.0 to 0.20 during the first 250 msec of the interval, was held at a plateau of 0.20 for 2000 msec, and then was decreased linearly from 0.20 to 0.0 during the last 250 msec. The period between stimulus intervals was 1500 msec. After each trial the observer indicated whether the standard velocity was faster or slower than the test velocity. The time between trials was typically about 3-4 sec. Each test velocity was presented 30 times in a random order with the constraint that no test velocity was presented the _n_th time before all the test velocities had been presented _n_-1 times. The velocity difference threshold ($\overline{DL}$) was derived from the function which related proportion of "slower" judgments to test velocity. The velocity $\overline{DL}$ was defined as one half the difference between the test velocities that yielded 75 and 25 percent "slower" judgments. In most cases, simple graphic interpolation (Kling and Riggs, 1971, pp. 25-26) was employed to determine the test velocities which corresponded to the 75 and 25 percent points. In cases when the function relating proportion of "slower" judgments to test velocity was nonmonotonic, or when none of the test velocities yielded proportions of "slower" judgments either greater than 75% or less than 25%, the proportions were first

converted to $\underline{z}$ scores and then a least-squares method (Kling and Riggs, 1971, pp. 27-29) was used to determine the test velocities corresponding to the 75 and 25 percent points.

Velocity $\overline{DL}$s were obtained for a number of standard velocities ranging from 0.3 to 20 Hz. The standard velocities were presented in a nonsystematic order. Measurements of $\overline{DL}$s were obtained with gratings of three different spatial frequencies in the following order: 4.8 c/deg, 0.6 c/deg, and 10 c/deg. The viewing distance for the 0.6-c/deg gratings was changed to 53 cm. At this distance the visual angles subtended by the parts of the stimulus display were approximately a factor of two greater than those described earlier.

The results of Experiment III are shown in Fig. 5. In each panel the velocity $\overline{DL}$ (expressed as a percent of the drift rate of the standard) is plotted as a function of the drift rate of the standard. Each panel gives the results obtained with a different spatial frequency. Each function in each panel describes the results for a different observer. A comparison of the vertical positions of the functions across the panels of the figure shows the effect of spatial frequency on the observer's ability to detect velocity differences. The vertical position of the functions shifts upward from the left to the right panel in the figure showing that the velocity $\overline{DL}$s increased as spatial frequency was increased from 0.6 to 10 c/deg. An analysis of variance of the data shown in Fig. 5 indicates that the main effect of spatial frequency is statistically significant. The $\underline{F}$ value for the main effect of spatial frequency was 98.1 with

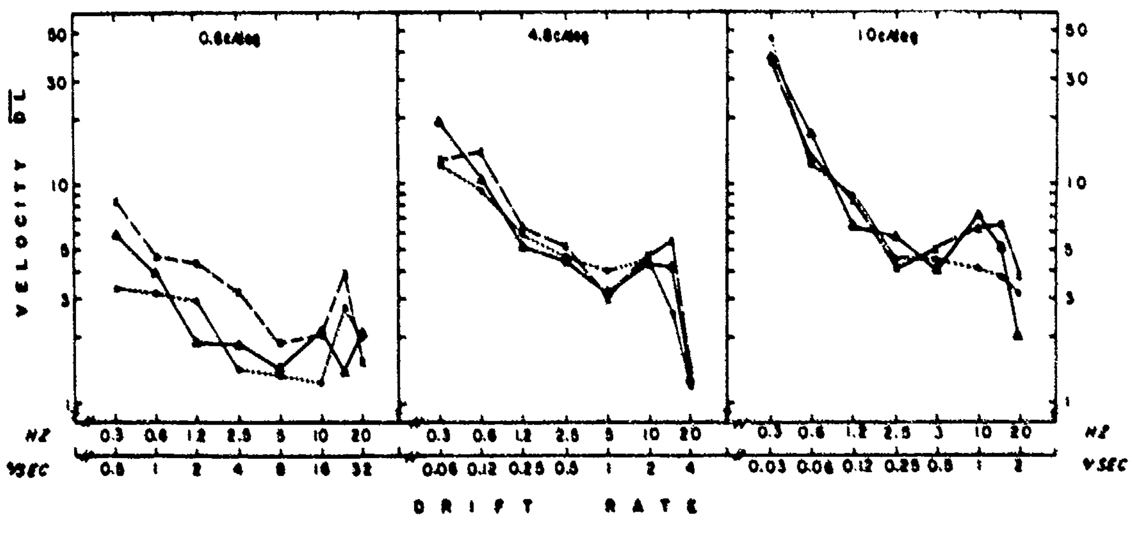

Fig. 5. Velocity difference thresholds ($\overline{DL}$s) as a function of the drift rate of sinusoidal gratings. Observer PLS: Xs; observer AJP: Δs; and observer KLS: Os.

2 and 4 degrees of freedom, and it is significant at the 0.002 level [$F(2,4) = 98.1$, $p < 0.002$]. The effect of spatial frequency on velocity $\overline{DL}$s is unlike its effect upon pattern recognition thresholds. Under comparable conditions, pattern recognition thresholds are lowest at an intermediate spatial frequency (around 5-6 c/deg) and increase at both higher and lower spatial frequencies (Kulikowski and Tolhurst, 1973). However, flicker thresholds like the velocity $\overline{DL}$s decrease monotonically as a function of spatial frequency (Kulikowski and Tolhurst, 1973).

Fig. 5 also shows that the velocity $\overline{DL}$ varies as a function of drift rate. The main effect of drift rate is statistically significant [$F(7,14) = 173.5$, $p < 0.0001$] and can be seen more clearly in Fig. 6 where the mean $\overline{DL}$s of the individual observers are plotted. Each function in Fig. 6 is the average of the three functions in one of the panels of Fig. 5, i.e., the average function for a particular spatial frequency. The relationship between the velocity $\overline{DL}$ and drift rate was essentially the same for all spatial frequencies, but the variations of the $\overline{DL}$ were more marked at higher spatial frequencies as indicated by a significant interaction of spatial frequency and drift rate [$F(14,28) = 22.0$, $p < 0.0001$]. Fig. 6 shows that velocity discrimination improved ($\overline{DL}$s decreased) as drift rate was increased from 0.5 to 5 Hz, grew slightly worse as drift rate was increased further to about 10-15 Hz, and then improved again for a drift rate of 20 Hz.

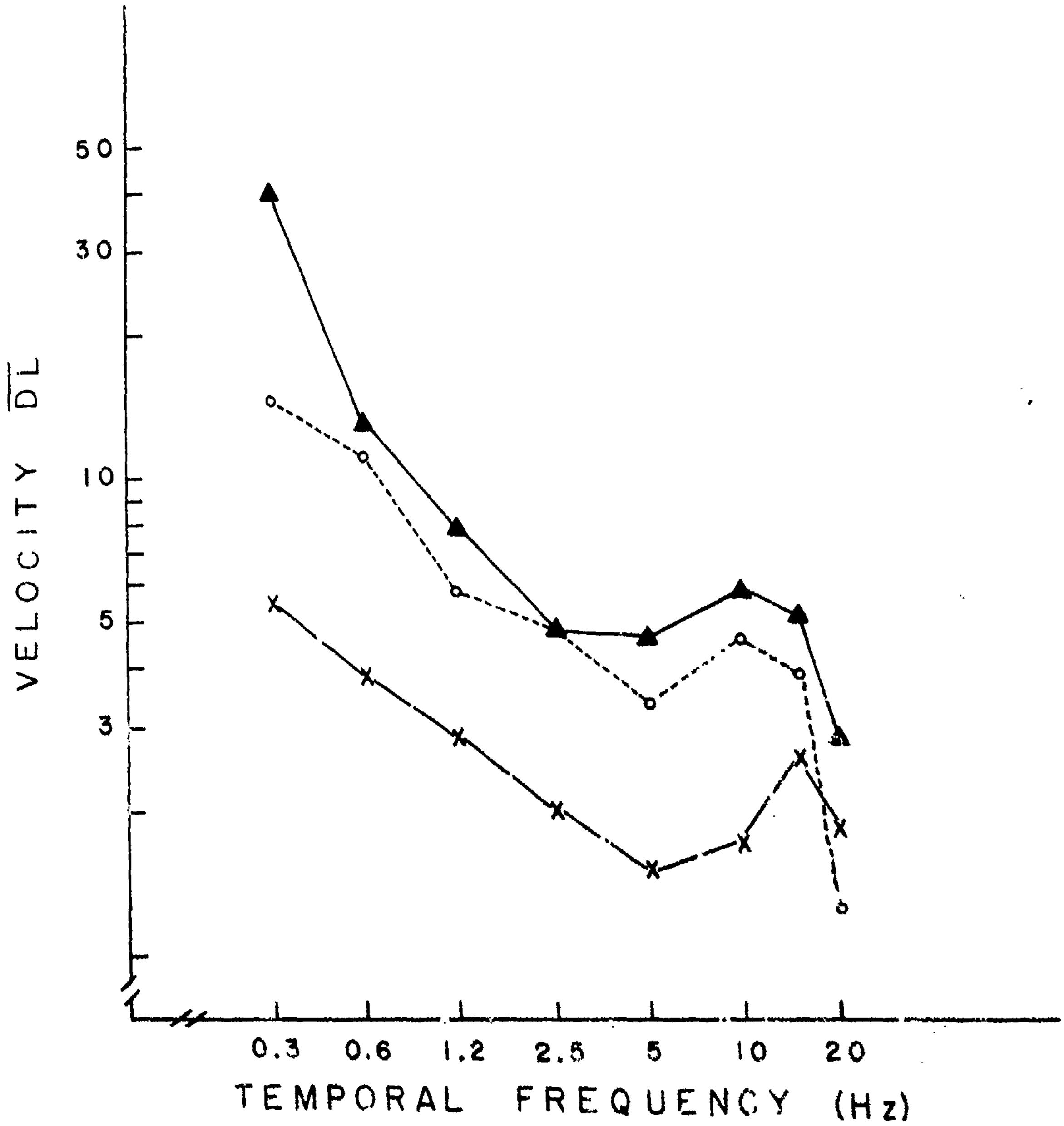

Fig. 6. Velocity difference thresholds ($\overline{DL}$s) as a function of the drift rate of sinusoidal gratings. Data depicted with Xs, open circles, and triangles were obtained with 0.6-, 4.8- and 10-c/deg gratings, respectively.

Comparisons of Temporal Frequency Response Functions of Experiments I, II, and III

Three different types of psychophysical measurements have been described above for specifying the temporal response properties of motion-sensitive channels: (1) measurements of contrast threshold sensitivity to drifting sinusoidal gratings (Experiment I); (2) measurements of the dependence of motion aftereffects on the temporal drift rate of sinusoidal gratings (Experiment II); and (3) measurements of velocity difference thresholds with sinusoidal gratings (Experiment III). In Fig. 7, temporal frequency response functions based on the three types of measurements are compared directly.

The function drawn with the solid line is a summary of the flicker threshold measurements of Fig. 1. The mean of the flicker thresholds of the individual observers (Fig. 1) was found at each temporal frequency. The mean thresholds were then normalized to give a maximum sensitivity (minimum threshold) of 0.0, and the normalized data are plotted in Fig. 7. The ordinate scale is inverted so that thresholds decrease (sensitivity increases) from the bottom to the top of the graph. As can be seen in the figure, flicker sensitivity is greatest at 2.5 Hz and decreases at both higher and lower temporal frequencies.

The function drawn with short dashes in Fig. 7 summarizes the aftereffect cancellation data represented by the open circles in Figs. 2-4. The mean of the null contrasts of the leftward moving grating (f_L) which were required by the individual observers to cancel the aftereffect generated by the rightward

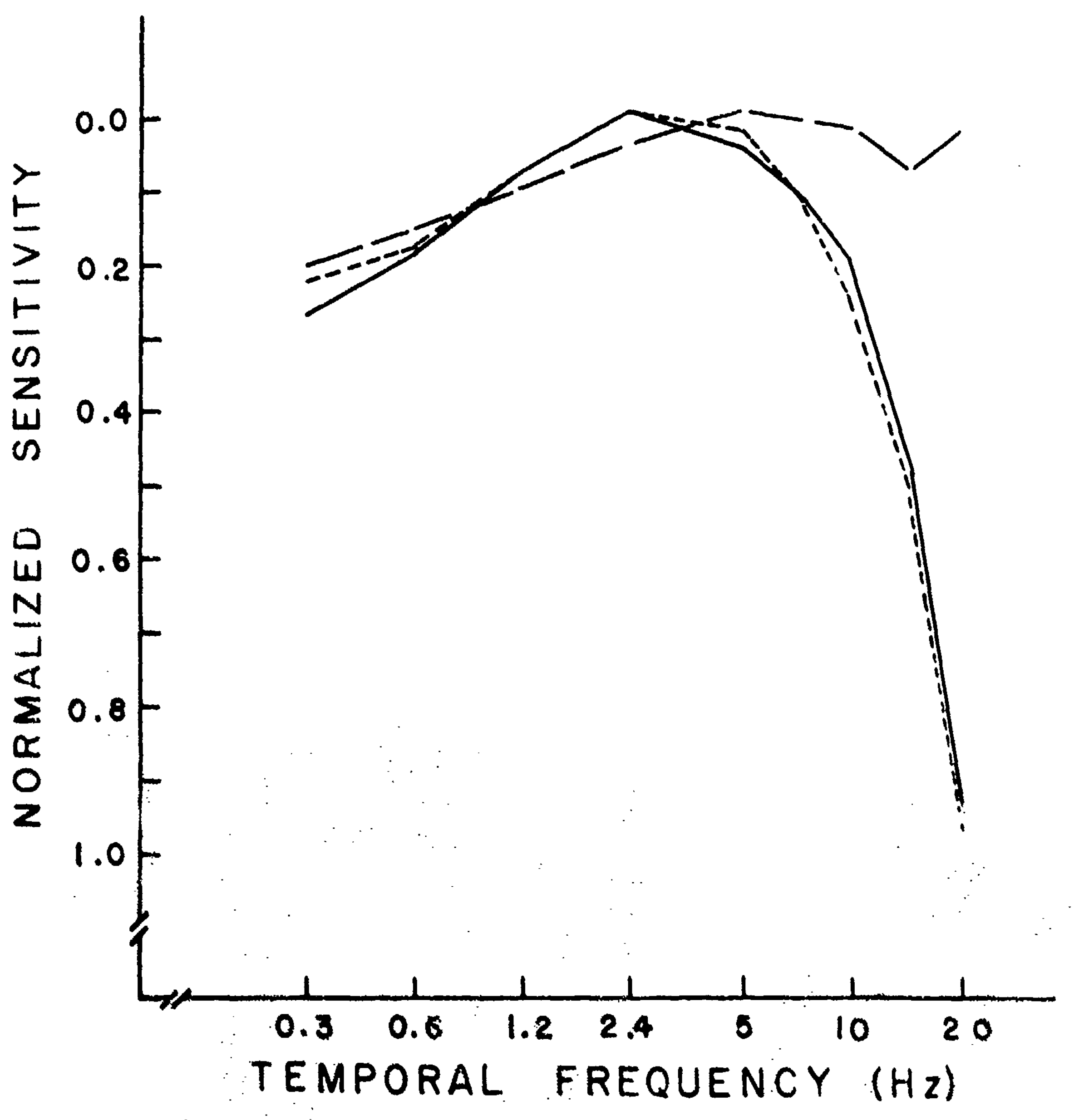

Fig. 7. Temporal frequency response functions derived from measurements of threshold flicker sensitivity (solid line), measurements of motion aftereffect cancellations (short dashes), and measurements of velocity difference thresholds (long dashes).

moving grating ($\underline{f}_R$) moving at 5 Hz* was computed for each drift rate of $\underline{f}_L$. The means were normalized to give a minimum null contrast of 0.0, and the normalized null contrasts are plotted in Fig. 7 as the short-dash function. As can be seen in Fig. 7, the null contrast function is identical to the flicker sensitivity function. The close correspondence suggests that common transient neural channels underlie threshold flicker sensitivity and the motion aftereffects. Moreover, we conclude that the response of transient channels is direction-specific. The null contrasts of Experiment II were contrasts required to cancel a directional aftereffect.

The function drawn with the long dashes in Fig. 7 was derived from the mean $\overline{DL}$ data represented by the Xs in Fig. 6. These $\overline{DL}$ data were obtained with the 0.6-c/deg grating, which was closest in spatial frequency to the 1-c/deg gratings used in Experiments I and II. The mean $\overline{DL}$s were converted to logarithms, and the logarithms were scaled such that (1) the minimum $\overline{DL}$ received a value of 0.0 and (2) the deviations of the scaled $\overline{DL}$ values from the average of the null contrast and flicker sensitivity data in the 0.3-to-2.5 Hz region were minimized. It is clear from Fig. 7 that velocity discrimination is unlike

*The shapes of functions obtained with a standard drift rate of 1.25 Hz are essentially the same as the shapes of the functions obtained with the standard drift rate of 5 Hz. Thus, it does not matter which of the two sets of data is used in the comparisons described below. However, the shapes of the functions obtained with the standard drift rate of 15 Hz were different. It may be that a drift rate of 15 Hz is near the temporal resolution limit of the motion channels which underlie the motion aftereffect. For this reason, the 15-Hz data were not used for comparisons in Fig. 7.

threshold flicker sensitivity and aftereffect sensitivity in that it does not exhibit a pronounced decline at high temporal frequencies. The basis of velocity discrimination at high temporal frequencies is presumably different from that of flicker sensitivity and the motion aftereffect. On the one hand, it is possible that velocity discriminations depend upon perceived differences in drift rate *per se* at low and intermediate frequencies, but upon other cues (e.g., changes in perceived contrast or blur) at high temporal frequencies. On the other hand, it is possible that velocity discriminations at high temporal frequencies involve a second movement-sensitive channel. The second channel may not contribute to threshold flicker sensitivity because it has a high contrast threshold, or to motion aftereffects because its response is not direction-specific. Evidence for two motion channels, one with a low- and one with a high-pass temporal frequency response, is described in the next section.

Research on Spatiotemporal Information Processing by Transient (Motion) Channels in the Human Visual System: Studies of Stroboscopic Movement with Multi-Element Displays

Element and Group Movement Sensations

Sensations of stroboscopic movement were produced by a cyclic alternation of two stimulus frames in a tachistoscope. Frame 1 contained 3 black dots (a, b, c) arranged in a horizontal row on a white background. Frame 2 contained 3 identical dots (d, e, f), also arranged horizontally but shifted to the right such that the positions of dots d and e of Frame 2 overlapped those of b and c respectively of Frame 1. With frame durations of 200 msec and an interval of approximately 40 msec between frames, the spatiotemporal display gave rise to a multistable percept: either the observer perceived a group of 3 dots moving *in toto* back and forth (*group movement*) or he perceived the overlapping dots of each frame as stationary and a third dot as moving back and forth from one end of the display to the other (*element movement*). We have been able to bring the multistable percept under stimulus control, i.e., to cause either the group movement sensation or the element movement sensation to predominate, by manipulating the duration of the interval between frames, the type of viewing (binocular or dichoptic), or the contrast of the stimulus frames. The results suggest that there are two systems or channels for generating movement signals in humans, each with different functional properties.

Apparatus

A 3-channel Gerbrands tachistoscope was used to present the two alternating stimulus frames. The viewing distance was 81 cm. At this distance each stimulus frame subtended a visual angle of 9^{o} horizontally and $6^{o}15'$ vertically. The diameter of the black dots in each stimulus frame was 40' with a center-to-center separation of 20' between a pair of adjacent dots. The luminance of the black dots was 0.15 mL; that of the white background 0.35 mL. In all the experimental conditions the duration of each stimulus frame was 200 msec. Only the interval between stimulus frames (ISI) was varied. The blank interval between stimulus frames was filled with a uniformly illuminated rectangle (0.10 mL), which was equal in its overall size to the stimulus frames (9^{o} X $6^{o}15'$) and was centered on the stimulus frames.

Experiment I: Dichoptic vs Binocular Viewing

Method

In the first experiment there were 12 different conditions resulting from the factorial combination of six interstimulus intervals (5, 10, 20, 30, 50, or 70 msec) and two types of viewing (either binocular or dichoptic). In the binocular viewing condition the observer viewed both stimulus frames with both eyes. In the dichoptic viewing condition, through the use of appropriately arranged polaroid filters, one stimulus frame was presented to the observer's right eye; the other stimulus frame, to the observer's left eye.* In both viewing conditions, the uniform

*Appropriate neutral density filters were used in the binocular condition to keep stimulus luminances equal to those during dichoptic viewing with polaroid filters.

rectangle presented during the interstimulus interval was seen by both eyes.

The dependent measure was the type of movement reported by the observer (either element or group) after he watched four cycles of one of the 12 experimental stimulus sequences.* While viewing each sequence, the observer was instructed to direct his gaze toward the center of the stimulus display (no fixation point was used) and at the same time to attend to (be aware of) the entire display. Each of the 12 experimental stimulus sequences was presented 10 times, following an order determined by block randomization. Eight observers participated in the experiment as part of a course requirement.

Results

The number of times that each observer reported seeing group movement in each of the 12 experimental conditions was converted to a percent. The pattern of results was the same for all observers. The percentage of group movement responses averaged across observers is shown as a function of ISI in Fig. 8. The solid function represents the data obtained in the binocular viewing condition; the dashed function, data obtained in the dichoptic viewing condition. As can be seen in the graph, the observers reported seeing group movement very infrequently (most

---

*When subjects arrived for the first session of the experiment they were shown (1) a stimulus sequence with an 80-msec ISI and binocular viewing and (2) a stimulus sequence with a 10-msec ISI and binocular viewing. After viewing each sequence, subjects were asked whether they perceived any movement of the dots, and if so, which dots moved and in what direction. All subjects spontaneously reported the movement sensation defined as group movement with the 80-msec ISI and the movement sensation defined as element movement with the 10-msec ISI.

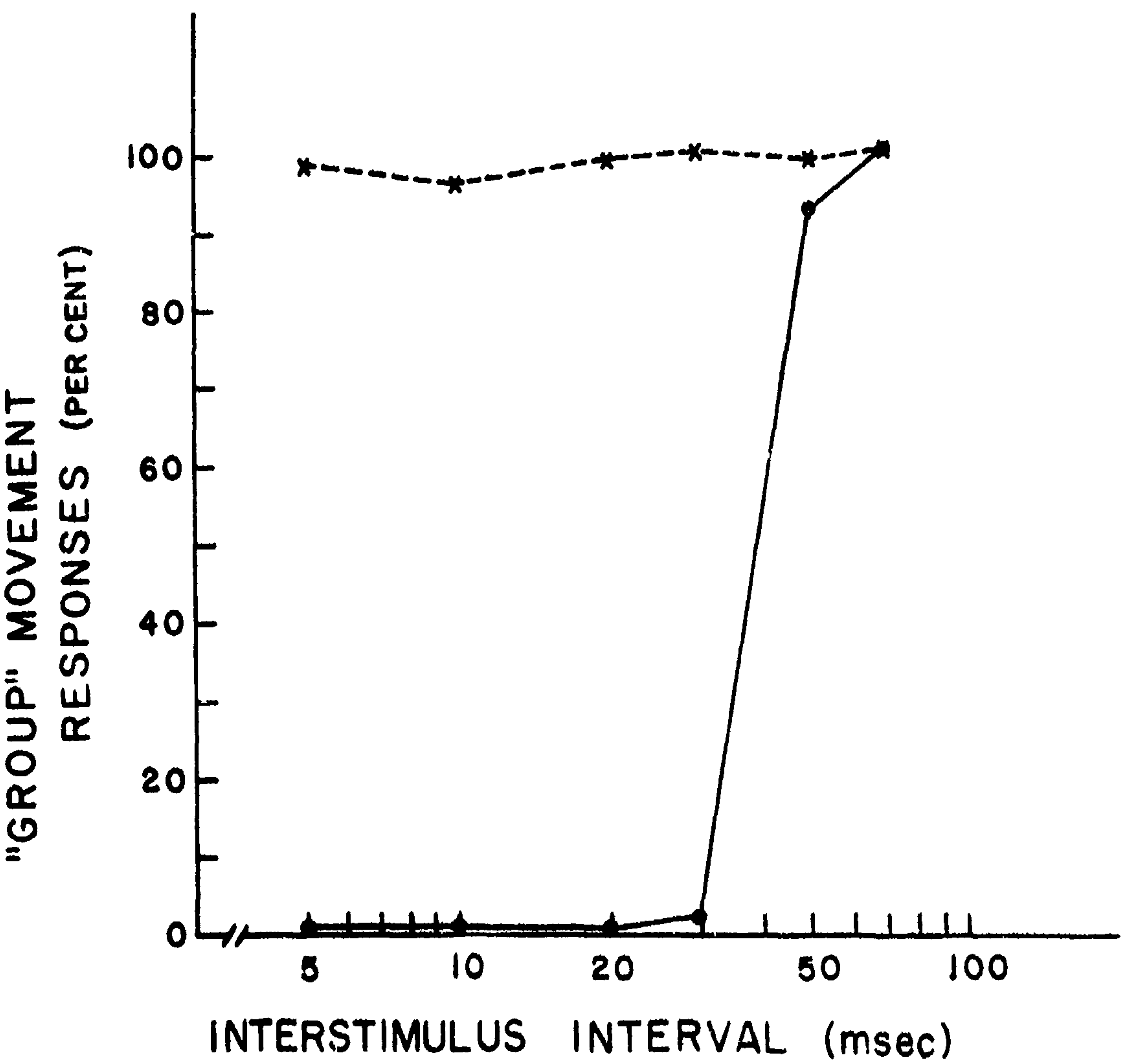

Fig. 8. Percentage of group movement responses as a function of the interstimulus interval between successive stimulus frames. Binocular viewing: ●s. Dichoptic viewing: Xs.

observers always saw element movement) at short interstimulus intervals (5, 10, 20, or 30 msec) in the binocular viewing condition. With binocular viewing and long interstimulus intervals (50 or 70 msec), the observers almost always reported group movement. In addition, the transition from the element movement sensation (few group movement responses) to the group movement sensation is fairly abrupt and takes place in the vicinity of 40 msec. In striking contrast to the results with binocular viewing, the mean percentage of group movement responses in the dichoptic viewing condition is equal to or near 100 at all interstimulus intervals.

Experiment II: Contrast Dependence

Method

The procedure used in the second experiment was identical to that used in the first experiment with the following exceptions. Two stimulus displays were used in the second experiment. In one condition (positive-positive), there were black dots on a white background in both stimulus frames. In the second condition (positive-negative), the dot-to-background contrast was reversed in the two stimulus frames. One frame contained black dots on a white background, and the second frame contained white dots on a black background. The luminance of the white areas in each frame was 0.25 mL; the luminance of the dark areas, 0.05 mL. The interstimulus interval was completely dark. The ISIs were 10, 20, 30, 50, 70, and 80 msec. The observers viewed the displays binocularly in all conditions. There were eight observers in the experiment.

Results

The mean percentage of group movement sensations reported in each condition is given in Fig. 9. The solid function shows the changes in the percentage of group movement sensations as a function of ISI in the positive-positive condition. These results are a replication of the binocular results of the first experiment. At short interstimulus intervals (10, 20, and 30 msec), the observers almost always saw element movement. At long interstimulus intervals (70 and 80 msec), the observers saw group movement on 90-100% of the presentations. By comparison, the mean percentage of group movement responses in the positive-negative condition is equal to or approaches 100 at all interstimulus intervals.

Two Transient Channels in Human Vision

The results of the stroboscopic movement experiments strongly suggest that there are two transient or motion channels with different functional properties.

The group movement sensation predominates (1) when long intervals (> 50 msec) intervene between the stimulus frames, (2) when the two stimulus patterns are presented dichoptically (i.e., Frame 1 is presented to one eye and Frame 2 to the other eye), <u>or</u> (3) when the contrast of the two stimulus frames is reversed (i.e., when Frame 1 contains black dots on a white background and Frame 2 contains white dots on a black background). These data indicate that the channel that generates the group movement signal is sluggish (i.e., has a relatively low-pass temporal frequency response), is located somewhere in the cortex

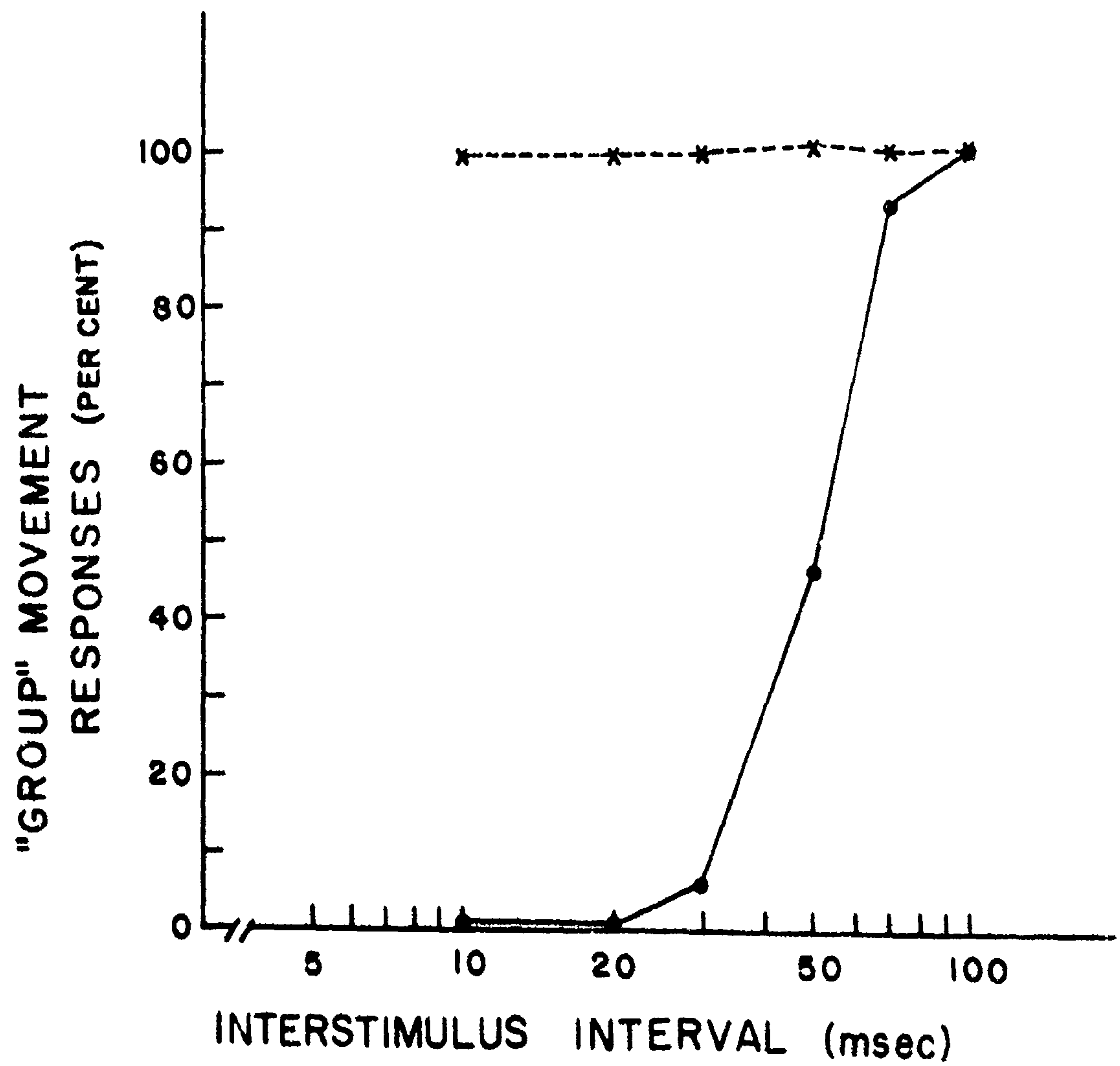

Fig. 9. Percentage of group movement responses as a function of the interstimulus interval between successive stimulus frames. Positive-positive condition: ●s. Positive-negative condition: Xs.

after the signals from the two eyes are combined, and is not affected by reversals of stimulus contrast. Some form or feature processing must occur at a stage prior to the site of generation of the group movement signal.

The element movement sensation is produced only if the interval between stimulus frames is short (< 50 msec), the successive stimulus frames are presented to the same eye, and the direction of stimulus contrast remains the same from one stimulus frame to the next. The data are consistent with the idea that the element movement sensation is produced by a spatiotemporal cross-correlation process; i.e., results from a neural computation which is functionally equivalent to a cross-correlation of the spatial intensity distributions of the two successive stimulus frames.

It is probable that the motion channel which generates the group movement signal is the channel whose frequency response was estimated from flicker sensitivity and motion aftereffect measurements. Group movement sensations predominated at ISIs greater than 50 msec (see solid functions in Figs. 8 and 9). Interstimulus intervals greater than 50 msec correspond to temporal frequencies less than 20 Hz, frequencies within the response range of the channel described by flicker sensitivity and motion aftereffect measurements. The discrepancy between the $\overline{DL}$ function and the flicker sensitivity and aftereffect functions at high temporal frequencies (see Fig. 7) might be attributed to a second high-pass transient channel, the channel that gives rise to the element movement sensation. The element

movement sensation predominated at ISIs less than 50 msec (see solid functions in Figs. 8 and 9). Interstimulus intervals less than 50 msec correspond to temporal frequencies greater than 20 Hz. The $\overline{DL}$ function in Fig. 7 may be a composite function determined by the independent contributions of the two motion channels isolated in the experiments on stroboscopic movement.

Research on the Recognition of Complex Real-Life Scenes

Studies of recognition memory for pictures typically involve two separate phases: (a) a study phase in which the stimuli are presented for viewing; (b) a test phase in which the stimuli are to be discriminated from other stimuli that have not previously been seen. These studies have shown that recognition became poorer as stimulus duration during the study phase was shortened (e.g., Loftus and Bell, 1975; Tversky and Sherman, 1975). Similarly, picture memory was poorer with a degraded image of the study stimulus, the test stimulus, or both, although degradation of one of the stimuli was more damaging than degradation of both (Dallett, Wilcox, and D'Andrea, 1968). The interpretation of these outcomes has been that the observer extracts more information from a visual stimulus as stimulus duration increases and that the presence of visual detail enhances recognition performance, as may the presence of verbal codes that supplement the pictorial information.

The method employed in the present experiments involves the two independent variables just described: (a) stimulus duration during the study phase of the experiment, and (b) whether the picture is in focus or blurred during the study phase, during the test phase, or both. (For convenience, these conditions will hereafter be referred to as conditions FU, UF, FF, and UU, the first letter describing the quality of the study image--focused or unfocused--and the second, the quality of the test image.) However, unlike past experiments, the present experiments derive

from a rather different theoretical rationale. Assume first the possibility that visual stimuli can be viewed as an aggregate of high and low spatial frequency information, the higher frequencies providing information about fine visual detail. The lower frequencies may also contribute to recognition memory by providing information about global blotches of light and shadow or other more general visual information. Assume second the possibility that separate visual channels exist for the analysis of high and low spatial frequency information and that these channels contribute to the formation of a memory code.

Given these basic assumptions, it is possible to state hypotheses about performance in the experimental paradigm to be employed. Depending upon the additional assumptions one makes about the order and rate of processing of information in different spatial frequency channels, different predictions can be generated about memory performance. The hypotheses and predictions offered below are not exhaustive but are intended to illustrate the usefulness of the paradigm. The intention is to pursue the research with these hypotheses in mind until the data inform us as to how they must be modified.

Suppose first that the low spatial frequency components of a stimulus scene were encoded first or more rapidly than the high frequencies. Then with extremely brief exposures of the stimulus during the study phase (study duration), the internal or coded representation of a visual scene would contain only low frequency information. For this reason, recognition of a blurred test image that contained only low frequency information (as in

condition FU) would be no better or worse than recognition of a sharply focused test image that contained both low and high frequency information (condition FF). With longer stimulus durations more information about low spatial frequencies will be encoded than with short study durations, and, in addition, high frequency information will also be encoded. Then, if the test image contained both low and high frequency information (condition FF), performance would be better than if only the low frequencies were present (condition FU). If the test image contained only low frequency information (condition FU), recognition performance would be no better than if only low frequency information were available during both the study and test phases (condition UU) because the high frequencies in the stored image would have no counterpart in the test stimulus.

In sum, this hypothesis predicts equivalent performance in all four focus conditions at brief study durations (FF = UU = FU = UF), improved performance with increasing study duration for all conditions, and differentially enhanced performance in condition FF at long study durations (FF > UU, FU, UF).

Conversely, suppose that the highest available spatial frequencies are processed faster than the lower spatial frequencies. If this were the case, performance in the FU condition would be relatively poor at short stimulus exposures. This is so because most of the information encoded during the study phase will be in the high spatial frequency range and the test stimulus would not contain the high frequency information necessary for recognition. As the stimulus duration is lengthened, performance in the

FU condition should more closely approach that of the UU and UF conditions as increasing amounts of low frequency information are processed during the study phase. It is difficult to predict how the FF condition would compare to the others. The improvement in this condition with increasing study duration, relative to the improvement in the remaining conditions, would depend upon the relative importance of the high and low frequency information in the target stimuli.

The final hypothesis to be considered here is that information is randomly sampled from the range of spatial frequencies available. If so, some proportion of the low frequency information and all of the high frequency information sampled during the study phase in condition FU will not be sampled during the test phase. Therefore, performance in this condition should be very poor, regardless of exposure duration. If, following Loftus and Bell (1975), we assume that the presence of visual detail enhances recognition, performance will be better and improve more rapidly with stimulus duration in condition FF than in any of the remaining three conditions. Performance in the UU and UF conditions should be between the FU and FF conditions (FF > UU = UF > FU).

Thus, different assumptions about the order and rate of processing of spatial frequency information generate different predictions about performance in the present experimental conditions. It is important to note, however, that these predictions rest on the assumption that the probability of correct recognition depends solely on the amount and kind of

information stored during the study phase and present during the test phase--forgetting, retrieval, and decision processes do not affect the conditions differentially. Also, regardless of whether visual or verbal codes are employed by the subject in performing his task, it is assumed that these codes are built up from information provided by independent spatial frequency channels.

Experiment I: Delayed Testing

Method

Subjects. The subjects were 40 paid volunteers, all students at Miami University.

Materials. The experimental scenes were chosen from three sets of vacation slides and were 35 mm black and white positive images. The long axis of each slide was horizontal. The slides were of outdoor scenes and were judged to be pictures of scenery rather than of a single object. Each experimental slide was classified as a member of some category (e.g., cities, houses, waterfalls). There were 10 categories in all and 6 slides from each category. Slides were back-projected by a Kodak Model 760H Carousel Auto-focus slide projector from a distance of 3.66 m.

Design. There were two independent variables: exposure duration (10 msec and 360 msec) during the study phase and degree of focus (focused or one-half turn defocused) during the study and test phases. Each subject served in only one of the four focus conditions: FF, FU, UF, and UU. Forty slides were presented during the study phase--four from each of the ten categories; of these four, two were shown for 10 msec and two for

360 msec. There were five orders of presentation of the stimuli for different subjects, but each slide always appeared in the same condition of degree of focus and study duration.

The test phase of the experiment consisted of a series of 40 slides evenly divided between targets (stimuli that had been presented in the study phase) and lures (items that had not been presented during the study phase). Only 20 of the 40 items used in the study phase were included in the test phase; 10 (one from each of the 10 categories) had been presented for 10 msec and 10 for 360 msec. During the test phase, the subject had 10 sec in which to examine each slide and to decide whether or not it was a target or a lure ("old" or "new"). Lures and targets were always presented in the same random order. Care was taken to insure that the last five items presented during the study phase were not among the first five presented during the test phase. The subject was informed at the beginning of the session whether the slides would be in- or out-of-focus during each phase of the experiment. Subjects were assigned to conditions by a randomized block method.

The subject sat at a distance of 0.91 m from the screen, and the visual angle subtended by the projected slide was 54^{o} horizontally and 38^{o} vertically. Table 2 shows the effect of the amount of defocusing on spatial frequency information. The tabled values are the fractions of a complete turn of the projector's focusing control needed to produce a threshold level of contrast in a particular projected sine wave grating. If a given fraction of a turn reduces a particular grating to threshold

contrast levels, we can assume that all higher spatial frequencies would also be effectively eliminated by the same fraction of a turn. Since the amount of defocusing used here reduced a sine wave grating of 0.2 c/deg to a threshold contrast level, it may be assumed that all higher spatial frequencies were also eliminated from the defocused slides.

Table 2

Amount of defocusing necessary to produce threshold contrast at several spatial frequencies

	Spatial frequencies (c/deg)				
	1.15	.58	.29	.2	.14
Turns	.17	.25	.33	.5	.87

Results

The initial analyses were done on the number of hits (a hit occurs when the subject identifies an "old" item as being "old") scored by each subject in each condition. For a given subject, this score could range from zero to ten, with chance level being five. The results are shown in Fig. 10. In general performance improved with increasing stimulus duration, and there was also an effect of focus condition. The effects were statistically significant, $\underline{F}(1,36) = 7.39$, $\underline{p} < .05$, and $\underline{F}(3,36) = 10.66$, $\underline{p} < .01$, respectively, as shown by a related measures analysis of variance. There was no significant interaction between the two variables, $\underline{F}(3,36) = 1.53$, $\underline{p} > .05$. According to a Newman-Keuls

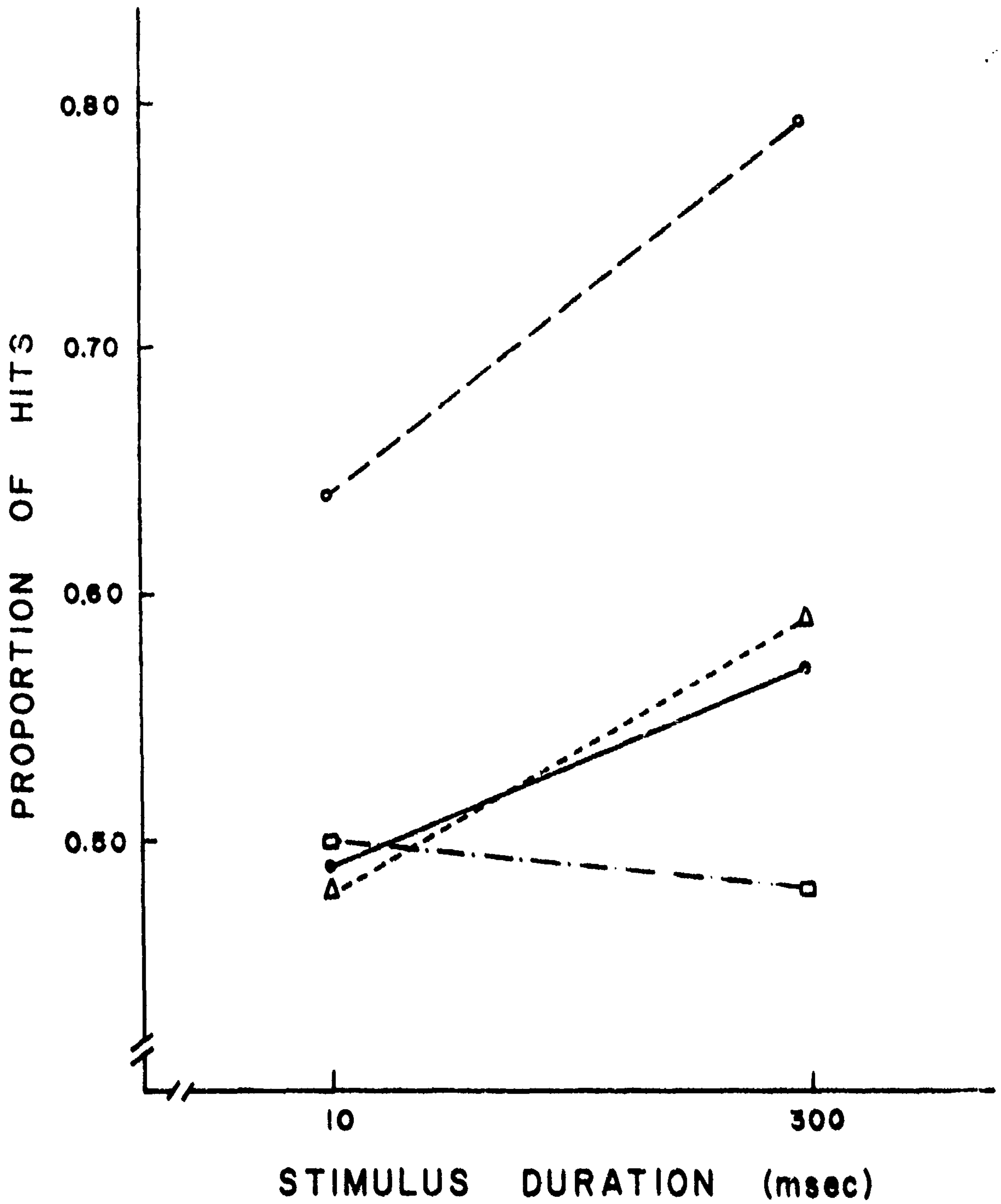

Fig. 10. Proportion of scenes presented during study phase which were correctly recognized during test phase as a function of stimulus duration during the study phase. Condition FF: Os; condition UU: Δs; condition UF: ●s; and condition FU: □s.

analysis, performance in the FF condition was superior ($p < .01$) to that in the other three, which did not differ among themselves ($p > .05$). This was true even when a separate analysis of variance was done on performance in the four focus conditions at the longer duration.

There were also differences among the focus conditions in the number of false alarms (saying "old" to a "new" item). A simple analysis of variance showed that the false alarms in the four focus conditions were significantly different, $F(3,36) = 13.8$, $p < .01$. A Newman-Keuls analysis showed that the FF group had the fewest false alarms ($p < .05$). The only other significant difference, according to the Newman-Keuls test, was that there were fewer false alarms ($p < .05$) in the UU group than in the UF group.

The number of false alarms indicates the extent of bias toward saying "old." The significant differences between the conditions in the number of false alarms indicates that this bias is not the same for all conditions. A simple way to demonstrate the effect of this differential bias is to transform the hit rates shown in Fig. 10 by subtracting the number of false alarms. These data are shown in Table 3. Since the lures erroneously

Table 3

Recognition (hit) rate corrected for false alarms

	Focus condition			
Stimulus duration	FF	UU	FU	UF
10 msec	40	8	11	-6
300 msec	54	18	9	2

recognized as "old" had not been shown during the study phase, half of the false alarms in each focus condition were allotted to each condition of study duration. From the transformed data, it is clear that performance is much better in the FF condition than in the others, which do not differ greatly among themselves. It is noteworthy that while the uncorrected data showed that condition UF was marginally superior to FU, the corrected data show poorest performance in condition UF. Table 3 suggests that performance in condition UF is below chance level at the 10 msec duration and only slightly above chance at the 300 msec duration.

Discussion

The data from this experiment do not provide a simple answer to the questions originally raised. There are large and statistically significant differences between conditions in the number of false alarms, and no analysis made solely in terms of processing priorities seems to account for all the data. It does seem, however, that the hypothesis that low frequency information is registered first may be rejected. The best evidence for this is the separation of the FF and UU conditions at the 10 msec exposure, which indicates that high spatial frequencies are registered even at this very short exposure. Each of the three original hypotheses predicted an interaction between stimulus duration and focus condition and a significant interaction was not obtained. One reason for this may be that none of the hypotheses are true. An alternative reason is that the data were distorted by the chance-level performance at the shorter exposure. It is clearly difficult for performance at the 10 msec exposure

to be any worse than it was in any condition but FF. One can only conjecture how the conditions would order themselves if the restriction were removed, and, for this reason, whether an interaction would be present.

Poor performance may also be the reason for the failure to find any differences among the UU, UF, and FU groups. The large number of false alarms in the UF group suggests that this group has less basis for successful recognition performance than the others, and, in effect achieves equivalent performance in hits at the price of a larger number of false recognitions. This is borne out by the transformed data of Table 3. This gives rise to the speculation that the UF group might perform less well than the others if the contribution of response bias were minimized.

As a subsidiary hypothesis, it was suggested above that high spatial frequency information might be of more use in recognition than is low spatial frequency information. If this is true, it would explain the superiority of the FF condition. However, it must be kept in mind that there is more information available in the FF condition where all the information in the picture is given at both study and test.

Because of low performance levels and the apparent intrusion of response bias in the recognition test data, Experiment I enables us to reject only the hypothesis that low spatial frequencies are scanned first or are processed more rapidly. The random sampling hypothesis and the hypothesis that high frequencies are processed first remain, and more complete information is needed before deciding between these two

possibilities. Furthermore, the suggestion that condition UF may be inferior to the others when decision factors are eliminated requires investigation, particularly in view of the results of Dallett, Wilcox, and D'Andrea (1968). While only a single, long stimulus duration was used, Dallett et al found that performance was best in condition FF and that condition UU was superior to that in both FU and UF conditions, which were comparable. Experiment II was designed to address these issues.

Experiment II: Immediate Forced-Choice Testing

A possible drawback of Experiment I was that the contribution of response bias was large. The false alarm data indicated that this was a possibility, and it is also true on a theoretical level. In attempting to understand the results of any recognition experiment, it is necessary to understand what the subject is doing when he makes his response. The Fourier-analysis model used here permits this in some detail for input or perceptual processes; however, the task becomes more difficult where output or response processes are concerned. Consider, for example, the problem of matching a test stimulus with items stored in memory in Experiment I. It is possible that the subject makes an exhaustive search of all the old items in memory, looking for a match. This seems unlikely in view of the large number of possible stored stimuli (cf. Anderson and Bower, 1973), and it is more probable that some restricted set is searched, although we do not know how the set might be specified. Thus, there is a need for a task that emphasizes perceptual, rather than memorial, processes, permits a simple conception of the response process,

and allows the subject to perform above the chance levels obtained in the first experiment.

To accomplish this, it was decided to use a procedure that involved a series of trials, each trial consisting of the study of a single picture and then an immediate test for recognition of the picture. A sequential forced-choice procedure was used in the test phase. It consisted of the successive presentation of a pair of pictures, one being the same as the target stimulus and the second being a completely unfamiliar exemplar of the same category. Thus, problems associated with a variable decision criterion were essentially eliminated; the subject merely had to identify the more familiar of the two test stimuli, a process that may also reduce the problems of retrieval.

Method

Materials. The materials were the 60 slides used in Experiment I--30 target stimuli and 30 lures. The sets of study and test slides were arranged in a random order. Each member of each test pair was presented for 5 sec, following as closely as possible upon the presentation of the target picture. The lure was presented first half the time. The degree of focus employed and the visual angle subtended by the slides were the same as in Experiment I, but the viewing distance was 1.83 m.

Design. As in Experiment I, the variables manipulated were stimulus duration (now either 5 or 200 msec) and the degree of focus during the study and the test phase. Independent groups of subjects were assigned to the eight conditions defined in this manner. The subjects were run in small groups, and groups were

assigned randomly to the various conditions. Study stimuli were presented by a Kodak Carousel, Model 800H, projector through an electronically controlled Gerbrands shutter. Test stimuli were presented by a Kodak Carousel projector, Model 850, whose defocused images were the same in character as those of the projector used in Experiment I.

Results

The major results are displayed in Fig. 11, where proportion of hits is shown for each experimental condition. Although performance was quite high over-all, there was a significant effect of focus condition, $\underline{F}(3,112) = 8.86$, $\underline{p} < .01$. There was, however, no significant effect of study duration, $\underline{F}(1,112) = 2.94$, $\underline{p} > .10$. A Newman-Keuls analysis revealed that the FF condition was superior to all others ($\underline{p} < .05$), that the UF condition was inferior to all others ($\underline{p} < .05$), and that the UU and FU conditions did not differ significantly from each other.

Discussion

The method of Experiment II improved performance considerably over the low levels obtained in Experiment I. The results of Experiment II also differ from those of the earlier study in that the subjects in condition UF performed less well than in the other conditions and in the lack of an effect of exposure duration. The latter change is of little interest as the difference does approach statistical significance, and it may be assumed (cf. Loftus and Bell, 1975; Tversky and Sherman, 1975) that a significant effect would be obtained with a greater difference of stimulus durations like that used in Experiment I.

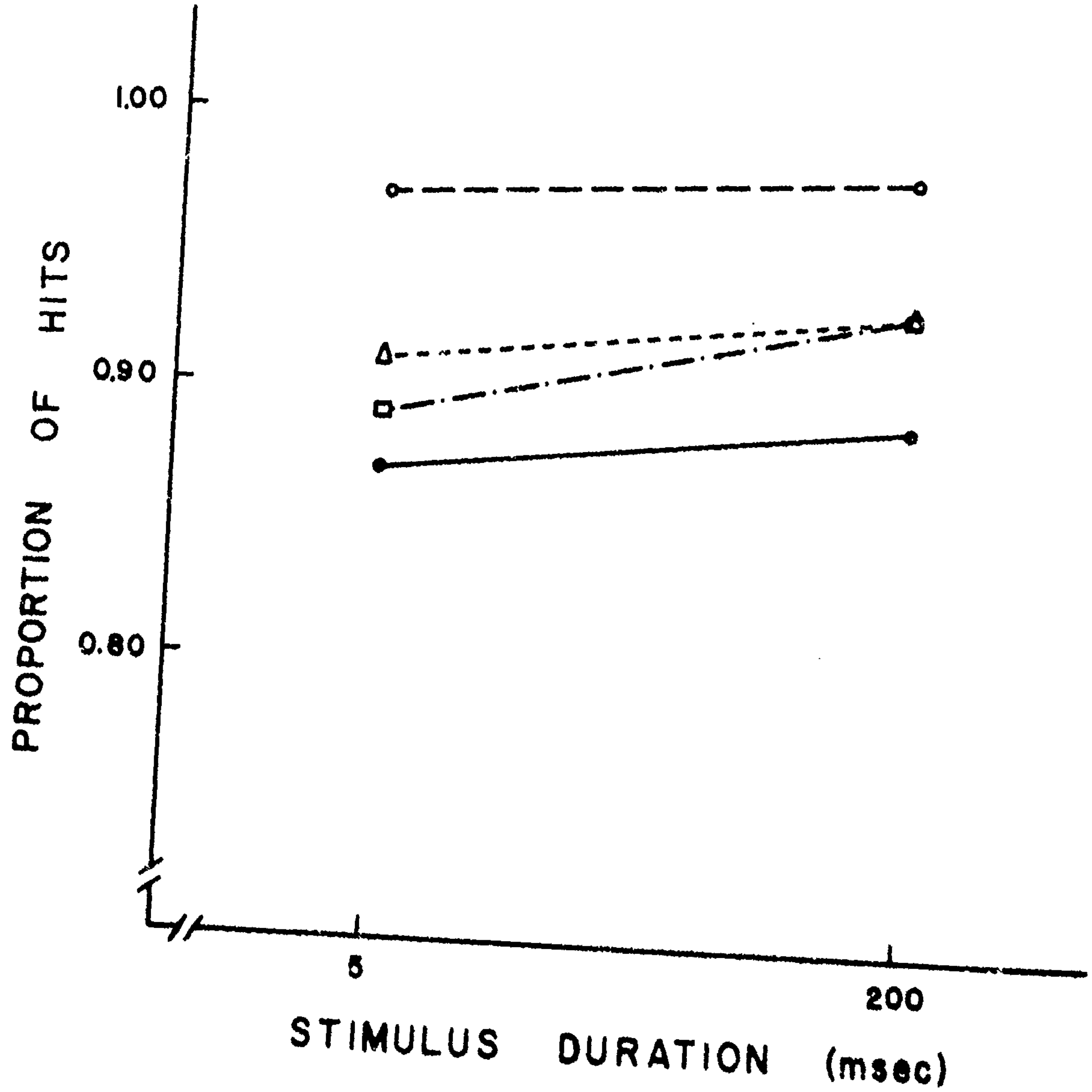

Fig. 11. Proportion of scenes presented during study phase which were correctly recognized during test phase as a function of stimulus duration during the study phase. Condition FF: Os; condition UU: Δs; condition UF: ●s; and condition FU: □s.

The near maximum performance levels tend to minimize the opportunity for observing an interaction between the variables. This is particularly true because the FF condition could not possibly show the relatively rapid increase with stimulus duration that was suggested in Experiment I. Nevertheless, a second failure to find an interaction between stimulus duration and degree of focus lends credence to the view that there would not be an interaction of the variables regardless of the absolute level of performance that is achieved.

The fact that the UF group performed less well than the other groups raises some interesting questions. Compared to the UU group, condition UF differs only in the presence of the high spatial frequency information during the test phase. The presence of this information somehow reduces the subject's ability to match the low frequency information in memory with that in the test slides. Possible causes of the impaired performance may be divided into three groups: First, there may be some perceptual difficulty, such as the masking of the low spatial frequency information by the high frequency information of the test slides (Pantle, 1974); in this way the activity in those channels of the visual system that signal the presence of low spatial frequency information would be reduced and recognition performance, in turn, would suffer. This is the type of explanation suggested by Harmon (1973) for an analogous finding in the recognition of block portraits. Second, the high spatial frequency information may dominate the subject's attention. While the subject is aware that the high spatial frequencies do not carry the information he

needs to make his decision, he may be unable to direct his attention to the low spatial frequency information. Third, and last, the difficulty for subjects in condition UF may have to do with the retrieval process. In condition UF the encoded stimulus contains only low frequency information, but the test stimulus, both low and high. Thus, if high frequency information in the test stimulus is dominant, it may interfere with the retrieval of the encoded image. The first two of these hypotheses are essentially reasons why the low frequency information in the test slide does not reach the memory system and emphasize perceptual or attentional mechanisms. The third suggests that memory may be unable to deal with this information when it is present.

In principle there is an attractive method for distinguishing between these two kinds of explanations for the performance in condition UF. Activity in high frequency channels inhibits activity in low frequency channels, but not vice versa (Pantle, 1974). Therefore, if there is a perceptual difficulty in supplying the appropriate cues in condition UF, we would not expect this mechanism to operate in an analogous experiment using only high frequency information in the study phase and both high and low frequency information during the test phase. This is so because the low spatial frequency components would not be expected to mask the high spatial frequency components. In other words, there should be no difference between performance in conditions using high frequency information in both study and test phases (condition HH) and a condition using high frequency information in the study phase and both high and low frequency information

in the test phase (condition HF). If the difficulty lies with attention, we would likewise expect no difference between the HH and HF conditions, as the relevant high frequency information should continue to dominate the low during the test phase. However, if the difficulty lies in retrieving a trace given a different retrieval cue, then performance in condition HH would be expected to be better than in condition HF.

In summary, the concern in these two experiments has been to elucidate the role of spatial frequency information in memory for pictures. In particular, we have dealt with three simple hypotheses about the order and rate in which such information is acquired. These do not exhaust the possible hypotheses, but they seem to form a suitable starting point. These hypotheses were examined in the context of two experiments in which exposure duration and degree of focus of the stimulus (at study and at test) were varied factorially. This method was successful in that it was possible to reject a simple hypothesis that the global, general information carried by the low spatial frequency components of a visual scene is acquired before other spatial frequency information. It was not, however, possible to reject either of the two remaining hypotheses. Further refinements in method may make this possible, but attention must also be given to exploring the reasons for the difference between the UU and UF conditions. The basic distinction made here was between memorial and perceptual difficulties, and a test of these two alternatives was proposed. Supplementary information about the acquisition of spatial frequency information in real-life scenes can be obtained

by studies of visual scanning. Techniques for studying eye scans are discussed in the next section.

Research on Visual Scanning of Complex Real-Life Scenes: Effect of Degree of Focus

Method

Visual processing of information in complex real-life scenes is related to the way in which the scenes are scanned (e.g., see Loftus and Bell, 1975). In order to determine whether there is a relationship between eye fixations and the spatial frequency content of local regions of complex scenes, we collected preliminary data on eye scans with the oculometer at Wright-Patterson Air Force Base. A free-viewing task was used in which observers were instructed to study each of 20 slides. The slides were selected from the set of 60 slides previously described in connection with the recognition studies of real-life scenes. The slides were presented on a screen 3.7 m in front of the observer. At this distance the slides subtended a visual angle of 22° in the horizontal direction and 16° in the vertical direction. Ten slides, one from each of the previously described categories, were presented in-focus and 10 slides were presented 5/12 of a turn out-of-focus with a Kodak Model 760H Carousel Auto-focus slide projector. Each slide was presented for 10 sec during which electrical signals corresponding to the observer's eye movements were recorded on magnetic tape. After each slide was presented, the observer scanned the edges of the slides in a designated order to provide calibration data. Records of eye movements were obtained for three observers. Two observers saw 10 slides in-focus followed by 10 slides out-of-focus. The other observer saw the in- and out-of-focus slides in reverse order.

Data Analysis

Positive 5 X 7 in prints were made of each of the 20 slides. The eye movement data in the form of electrical signals on magnetic tape were low-pass filtered with a high-frequency cutoff of 10 Hz. The filtered signals were amplified and fed into an X-Y plotter to reconstruct the scanpaths of an observer. The scale of the X-Y plot was adjusted to match the 5 X 7 in prints of the slides. The scanpaths were then used as overlays on the prints to determine the order of points fixated in each scene. Further analyses will correlate eye fixations with the frequency content of local regions of the scenes. Correlations between eye fixations and the local spatial frequency content of scenes can provide useful information about the search strategies employed by observers in different perceptual tasks.

References

Anderson, J. and Bower, G. _Human Associative Memory_. New York: Winston and Sons, 1973.

Dallett, K., Wilcox, S. G., and D'Andrea, L. "Picture memory experiments." _Journal of Experimental Psychology_, 1968, 76, 318-326.

Harmon, L. D. "The recognition of faces." _Scientific American_, 1973, 229, 70-83.

Kling, J. W. and Riggs, L. A. (Eds.) _Woodworth and Schlosberg's Experimental Psychology_. (3rd ed.) New York: Holt, Rinehart and Winston, Inc., 1971.

Kulikowski, J. J. and Tolhurst, D. J. "Psychophysical evidence for sustained and transient detectors in human vision." _J. Physiol_., 1973, 232, 149-162.

Loftus, G. R. and Bell, S. M. "Two types of information in picture memory." _Journal of Experimental Psychology_, 1975, 104, 103-111.

Pantle, A. "Visual effects of sinusoidal components of complex gratings: Independent or additive?" _Vis. Res_., 1973, 13, 2195-2204.

Pantle, A. _Visual Information Processing of Complex Imagery_. Aerospace Medical Research Laboratory Technical Report -74-43, Wright-Patterson Air Force Base, Ohio, AD785366, July, 1974.

Pellegrino, J. W., Siegel, A. W., and Ohawan, M. "Short-term retention of pictures and words: Evidence for dual code systems." Journal of Experimental Psychology: Human Learning and Memory, 1975, 104, 95-102.

Tversky, B. and Sherman, T. "Picture memory improves with on time and off time." Journal of Experimental Psychology, 1975, 104, 114-118.